

Citation for published version:

Snider, C, McAlpine, H, Gopsill, J, Jones, S, Shi, L & Hicks, B 2015, It's Not Personal: Can Logbooks provide insights into engineering projects? in C Weber, S Husung, M Cantamessa, G Cascini, D Marjanovic & S Venkataraman (eds), *DS 80-10 Proceedings of the 20th International Conference on Engineering Design (ICED 15)* Vol 10. vol. 10, IECD, Design Society, pp. 219-230.

Publication date:
2015

Document Version
Peer reviewed version

[Link to publication](#)

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 INTRODUCTION

Today's machines and products are so advanced in terms of their materials, form, construction, control and drive systems that they require expertise and resources that stretch even the world's largest organisations. Failure to understand and control the myriad activities that typify engineering projects leads to at worst complete failure (with large associated financial and reputational penalties) and at best, cost overruns, delivery setbacks, leakage of Intellectual Property (IP), exposure to unnecessary risk, and difficulties capturing lessons learned, decisions and rationale.

The high price for 'getting it wrong' has led to significant research and subsequent guidance on what a good project (and good project management) looks like (see OGC, 2009; Pinto & Slevin, 1987). However, applicability of generic performance and monitoring principles is often problematic. For example, generic high-level understanding states that high communication rates are good (Katz 1982), but does not consider the individual context of each project where this might not be the case; such as in those tasks better completed by individuals (see Dunnette, Campbell, & Jaastad, 1963), or when a high individual rate of work is required to meet deadlines or cost restrictions.. We also contend that much of the existing guidance and 'best practice' is often too general to be useful. The wider research programme (of which this research is a part) takes a different approach to understanding engineering projects, with the goal of generating high level understanding from low-level engineering project data.

So far, this research programme has analysed information typically generated in large quantities during projects, including email, reports, CAD and other electronic files. They have been analysed with a range of techniques, including sentiment analysis, various types of cluster and content analysis. The early indications are that there is indeed a lot of insight and understanding to be gained from analysing the records of projects, both in terms of learning to benefit future projects, but also providing real-time or near real-time actionable information to engineers and project managers, alerting them to, for example, sudden changes in sentiment in relation to a particular sub-assembly, the potential creation of IP, or the lack of treatment of risk in a project.

To date, all the information types tackled have been computer tractable, either capable of being parsed directly, or via monitoring or transformation using Optical Character Recognition (OCR). We recognise that despite an increasing number of types of digital information used in engineering projects, there still exists a large proportion of non-digital, often less formal records. In particular, engineering logbooks are well recognised as important and unique source of information not found elsewhere (McAlpine et al. 2006). In the case of logbooks, their largely paper-based format, mixed media content and personal nature makes the automated extraction and application of analysis extremely challenging. Manual extraction or coding of informal information such as logbooks clearly then carries with it a large time and therefore cost overhead.

We also observe that much existing research concerning engineering logbooks has been conducted from the perspective of understanding their use and content with a view to digitising the 'logbooking' process. The objective is often to support the logbook user and to provide benefits to the organisation such as compliance with record-keeping and presumed rationale capture (Gross & Do 1996; Oehlberg et al. 2009; McAlpine et al. 2006). However, the focus of this paper is different: Instead of asking 'how can we make this digital?' we are seeking to answer the question 'what do we need to capture from logbooks to provide actionable insights to engineers and project managers?' We are not looking to change or better support logbook-keeping practice, but rather to gather information to inform higher level understanding of projects from the information contained within them.

The aim of this research is therefore to discover whether paper-based engineering logbooks contain information from which project managers and engineers can potentially derive useful, actionable insight and ultimately a better understanding of project performance.

To do this, we have manually applied two existing coding schemes to a set of seven paper-based engineering logbooks. The coding schemes are detailed in section 2, before selected results are presented in section 3. Section 4 then discusses the implications of the findings for providing insights into - and understanding of - engineering projects. Whilst this preliminary study shows that logbooks could indeed be a rich source of potentially valuable information to provide insight and aid project understanding, section 4.1 discusses the barriers to achieving this, and some practical steps that could be taken to make the information more amenable to automatic analysis.

2 METHODOLOGY

To support the aim of discovering whether paper-based engineering logbooks contain information from which project managers and engineers can potentially derive useful, actionable insights into project performance, a set of three engineering logbooks have been manually coded. One schema focusses on the content of the logbook, and one focusses more on the context of the tasks undertaken. Both schemas were tested for inter-coded reliability with results presented below, and process described in their referenced texts.

All logbooks were taken with permission following the completion of a 22-week final-year undergraduate design project at the University of Bath, during which each engineer developed a product from initial brief to proof-of-principle working prototype. Product type varied within each project, including a submarine thrust system, a sonar mapping system, and a head restraint for cerebral palsy patients, but the process followed and coverage of the typical design process did not. As the type of product is not the primary determinant of the types of design and process followed by engineers (Snider et al. 2014), this gives the opportunity to compare processes and learnings from each engineer directly.

In each case, the logbooks were divided into discrete 'tasks' - chunks of information that related to the engineer working on a particular topic. In reality, the tasks closely followed the logbook entries, which were usually clearly defined with a title and date.

Due to space constraints, each coding scheme used in this paper is described by its relevant components only in the following sections; each has been used extensively in other research with validation and results presented elsewhere and referenced in the text.

2.1 Coding for Content

The coding scheme adopted for this research had originally been developed by Wasiak et al. (2008) to classify engineering email content. It was created from an extensive review of the literature on classifying information in various fields, including engineering design, sociology and organisational behaviour. The categories cover i) How the information is presented (e.g. whether it is a written note, sketch, calculation etc.), ii) What it is about (whether it is product or process-related and what aspects, such as product performance or materials, or project planning, risk or cost it concerns) and finally iii) Why it is being created (in terms of problem solving activities such as exploring or making a decision and the intent of the communication, such as informing, exploring or debating).

The studied logbooks each contained between 156 and 219 discrete tasks, with all content of the logbooks being coded for each category of the scheme. Intercoder reliability for this scheme is reported elsewhere (Wasiak et al. 2010), with coefficients between $\kappa = 0.71$ and $\kappa = 0.87$ (Cohen 1960) for each section.

2.2 Coding for Context

The coding scheme adopted for this section of the research is designed to identify at an abstract level the types of task that engineers complete throughout their design process, with a particular focus on tasks with a creative manner. Designed specifically for use with logbooks with a combination of inductive and deductive approaches, this coding scheme has been described in detail, demonstrated and validated in other research (Snider et al. 2013; Snider et al. 2014).

The scheme itself identifies creative behaviour through evidence of expansion within tasks, described as variation away from well-defined schemas and the exploration of options, as opposed to following well-defined and structured methodologies. Additionally, the scheme studies task types through focus of the designer either on the development of information for use within the design, or on the physical or virtual design itself as an entity itself. These two categories allow understanding both of what a designer is trying to achieve within the wider design process context, and the means of their achieving it in terms of creative behaviour. Being highly latent in nature the coding scheme requires significant interpretation from a human coder, but achieves inter-coder reliability of $\alpha = 0.768$ (Hayes & Krippendorff 2007), suitable for this form of exploratory work (Blessing & Chakrabarti 2009; Klenke 2008).

Instead of focusing on the explicit content of any entries made by an engineer, this scheme looks at the types of tasks that they were completing, the nature of those tasks, and how each fits into the wider process. As a result, it informs less about the manifest and details of the work being completed,

and more about the higher-level patterns and trends in activity influenced by the skills and qualities of the engineer, type of design being completed, and stage within the design process.

2.3 Interpreting the Data

The purpose of the interpretations presented in the following section are to illustrate the time-based learning that can be derived from engineers logbooks at a project level, allowing the generation of actionable information for an informed manager or interested engineer.

To contextualise the findings within the scope of potentially useful information, each is compared to what are termed project *features*, those aspects of an engineering project whose state are important to ensure overall performance and the development of a quality output. While these features are developed and presented elsewhere (Snider et al. 2015), in this paper each finding is presented only in context of the relevant literature to which they relate – in each case, the finding informs about the state of an aspect of the project that in part determines overall performance and output.

By understanding these project features in detail through the types of analysis presented here, the manager may make informed, evidenced-based decisions on intervention within their projects. Consequently, they may have better control over the performance achieved within projects, and a means to monitor any interventions that they may make.

3 RESULTS

Results from each coding scheme are presented here separately for both brevity and clarity. In practice, substantial scope exists for combinatory analysis of such schemes, particularly in comparison of such manifest and latent schemas (Cash & Snider 2014). The results here are presented in two streams, the first focused on producing “baselines”, or understanding of typical occurrences within projects against which current projects can be compared; the second focused on interpretations that can be made directly from patterns observed in data. Although the content scheme is used for the former and the context scheme is used for the latter, both are equally applicable to each role.

3.1 Establishing Baselines

The results for the content coding scheme can be studied from a time-based perspective according to the main categories by which the scheme coded: what the content is about, and why it is being created. From a project management perspective, particularly due to the time-based analysis, each of these has potential to provide insight into an engineering project as it progresses.

Figure 1 shows what the content of each logbook was about for each of the three participants in tandem. Each row indicates the content of a specific task (non-mutually exclusive), with a black marking indicating the appearance of that content type in each task. Content can be roughly decomposed into project-based (which is of a more managerial nature e.g. planning activities) and product-based (of a more product-based nature e.g. exploring product performance), as labelled in each sub-figure. As all projects were of the same timescale and approximate types of working were similar within the timescale for each, they can be generically compared by relative proportions, despite the varying number of tasks completed by each engineer. In all figures within this paper, larger grey shading indicates general stage transition areas, where work shifted from mainly concept type, to mainly embodiment type, to mainly detail type.

In each sub-figure, there is a distinct predominance of *Product-Feature* type content (that developing the physical features of the output), followed by a similarly high-proportion of *Project-Planning* type content (that planning future project activities and directions). Within each sub-figure, a slight tendency for planning to occur earlier in the process can be seen, followed by a higher amount of *Product-Performance* in middle-stages iterated through *Product-Feature* content towards the end. Each of these trends makes some logical sense in terms of a typical design process – in early stages an engineer needs to plan what they are going to develop, and in later stages of more detailed design, they need to evaluate the performance of their chosen solutions (Pahl & Beitz 1984; Pugh 1990). Throughout the process, as their goal is the development of the product, all content is iterated against development of the features of the final product.

Rather than describing distinct patterns in this data, for the project manager such analysis gives value directly in the information generated. The knowledge of what content is being developed through the process gives an understanding of the focus of engineers which can, at the most basic

level, be compared against the manager's own expectation to imply whether the project is *normal* or *atypical*. For example, in expecting a higher focus on planning in early stages and high iteration between features and other content types, upon seeing either of these features (or noticing their absence) a manager can gain an informed understanding of whether the project is progressing as they would expect it to given their implicit understanding of its content and context. For example, the absence of consideration of product cost in the early stages of the design where much of the cost of a product is set, could be a warning sign. Similarly, if a high amount of work is completed in a different area that is not expected (such as concentrated work on materials or manufacturing when not usual) the manager can intervene, should they feel it is needed.

In contrast to distinct and algorithmic patterns within the data, the time-wise understanding of content generated throughout the project gives the manager the means to understand progression and compare against expectation, hence giving the base-level understanding necessary to intervene when required/appropriate.

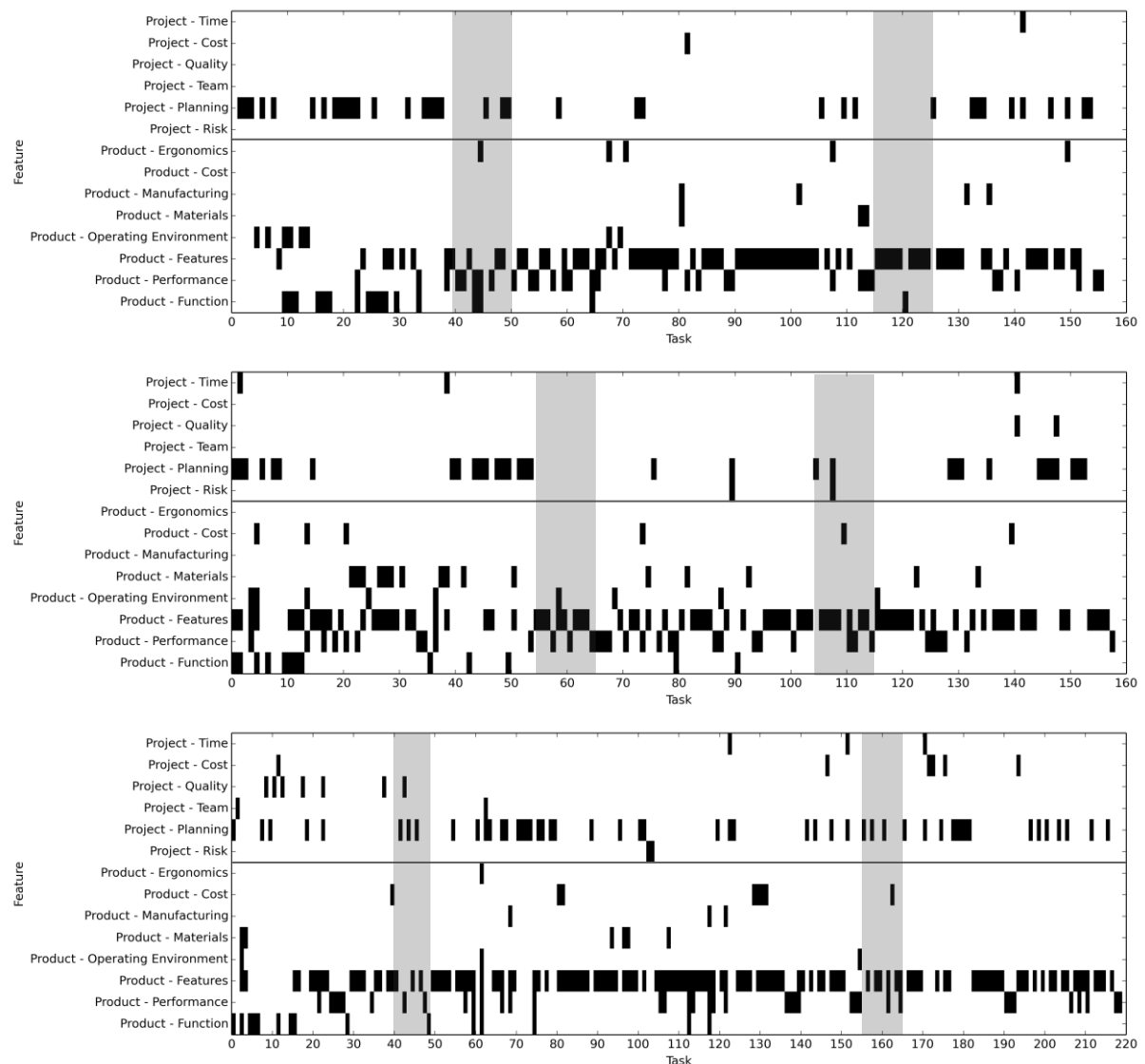


Figure 1: "What" the engineers were doing (grey areas indicate approx. stage boundaries)

In a similar way to Figure 1, Figure 2 describes occurrences evident in the markings of engineers in each task. In this case, however, coding is for the *why* element of the content coding scheme, describing in this case the type of problem solving completed by the engineer at each point in time.

In each subfigure there is a clear majority of *problem solving* type activity displayed. This encompasses searching, gathering and exploring solutions. This is logical given the nature of design as a problem solving process (Simon 1973; Dorst 2003). For engineers A and C, there is a predominance

of *Goal Setting* type activities earlier in the process and then at interspersed intervals throughout; while *Engineer B* displays less *Goal Setting* but more *Evaluating* type activities throughout. In all cases, there is little evidence of either *Constraining* or *Decision Making* within the logbooks.

Such apparent individual differences as seen here between *Engineers A and C* and *Engineer B* illustrates the importance of interpretation of data by an informed observer. While with data it is possible to study the cause of such differences in more detail, in practice, within industry such a process may not be feasible. The analysis and understanding generated by any algorithmic methods of analysis should not require excessive additional and targeted work by those within the company for interpretation. However, by presenting such figures as seen to a manager their understanding of the project context may allow their own interpretation without further detailed analysis. For example, should each engineer work within a team with the data presented to a project manager, they will understand their workers to the extent that they can determine the patterns as normal or a-typical, based on their understanding of the typical working of each person and the project. Similarly, should such differences be caused by process variations, their understanding of the process being completed will allow them to recognise, interpret and take appropriate action.

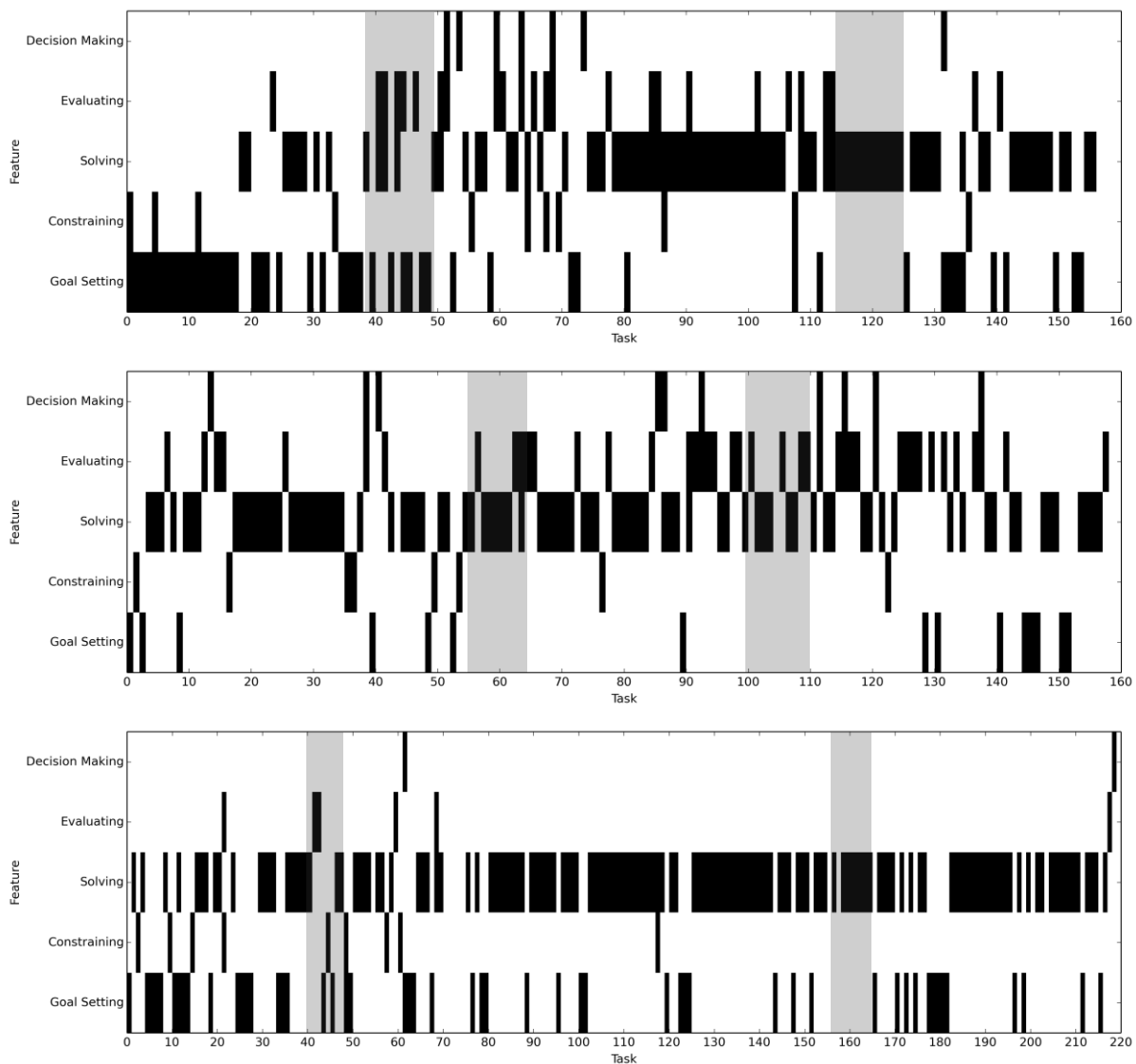


Figure 2: "Why" the engineers were doing each task

From a managerial perspective, such data as Figure 2 informs about the type of work being completed by each worker, again allowing the manager to compare against their idea of *normal*. Given that a design process is a *problem-solving* activity, the manager may expect to see predominance at each point in the process. Given that the manner of progression is less clear in conceptual design and many process are stage-gated or passed between teams, the manager may expect to see more *Goal Setting* at

certain points, and given that the manager should be aware of the general content and working patterns of their team, they may expect to see more or less *Evaluating* behaviour throughout the process. Such analysis and representations as seen in Figure 2 therefore demonstrate utility of such analysis from the project management perspective – in seeing with some detail what sort of activity is being completed at each stage, a manager can intervene when typical patterns do not occur, should they deem them to be required in their context.

3.2 Patterns in Data

The results from the context coding scheme presented here focus on the creativity and novelty evident in the design process of each engineer, looking more at the specific patterns that can be seen and information that can be inferred from them. Due to space constraints the coding scheme used to generate these results is not presented here. The examples given describe in brief the concepts that the coding scheme describe and the potential actions of the manager, but for a detailed understanding of the coding scheme and implications of the results, the referenced texts should be reviewed.

Figure 3 shows the occurrence of creative behaviour in each task for each engineer, where an upward stroke indicates a creative task, and a downward stroke indicates a non-creative task. Grey shaded areas indicate general transitions between process stages.

The differences between Engineers A and B, and Engineer C are very evident. In conceptual stages of the design process, while both A and B are reasonably creative in their tasks with a general upwards trend, Engineer C is primarily non-creative. This may be an area in which a manager wishes to act; a lack of creativity in early stages when compared to peers may, depending on the product and circumstances, be a negative that should be rectified. Through both embodiment and detail stages all engineers demonstrate similar behaviour with a slight non-creative trend through embodiment, and a steeper decline in creative behaviour in detail. This pattern may be an indicator of normal, against which a manager may compare their own expectations and other projects. Should they see more or less creative behaviour in a project, it may be indicative of either positive or negative behaviour that could be encouraged or prevented.

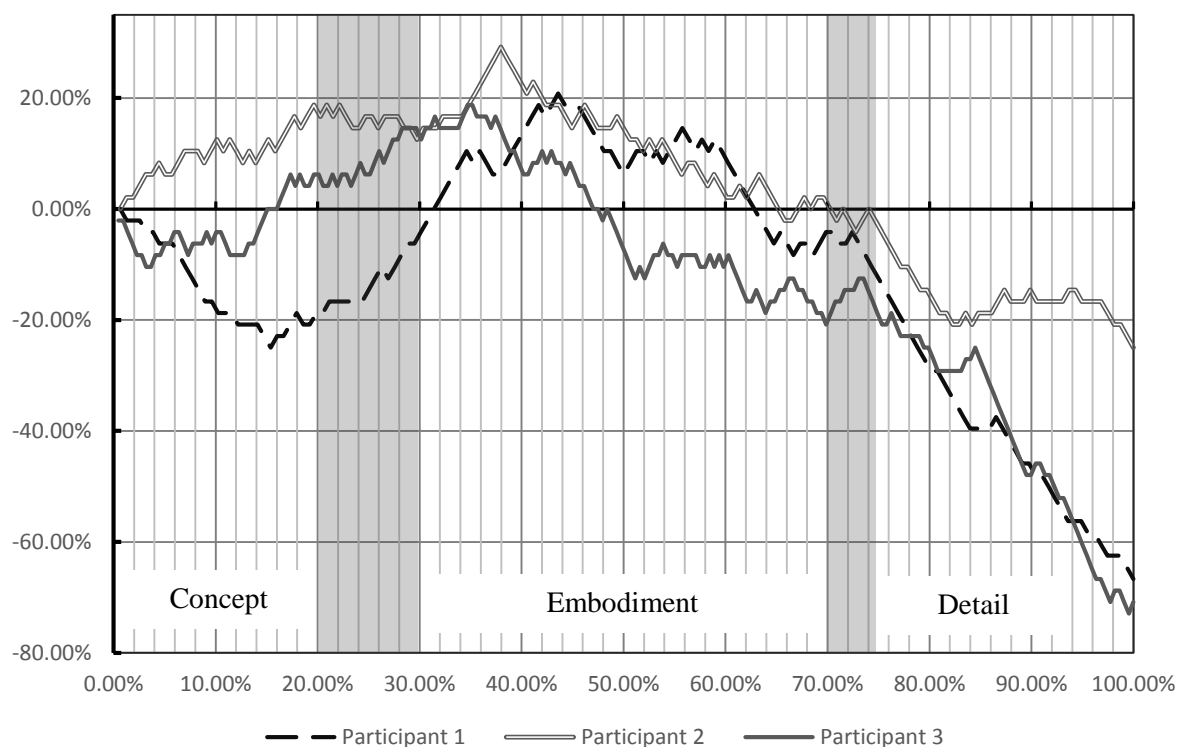


Figure 3: Occurrence of creative behaviour through the design process (grey areas indicate approx. stage boundaries)

For example, in the detail stage, Engineer B shows a sudden increase in creative behaviour towards the end of their process. This may be an area in which a manager should investigate. As creative

behaviour is explorative (Guilford 1956), an increase may indicate an issue that the engineer is attempting to solve which, particularly at such a late point in the overall process, could be critical to observe and manage. Similarly, as creative behaviour encourages novelty in solution, such a display of creative behaviour at the end of the process could indicate potentially higher novelty in output, which may cause issues in interface with related sub-systems.

Thus from a managerial view, an indicator of novelty has potential to give an understanding of process progression, and the occurrence of a-typical activity within the design process which may require observation or action on the part of the manager.

Figure 4 also shows creative behaviour, but in tandem with the types of task completed by each engineer. Divided into tasks that develop the physical or virtual design itself (upper in each subfigure, termed *application-type*) and tasks that develop or increase the information content of the design process (lower in each subfigure, termed *information-type*), Figure 4 shows patterns in the types of activities that are completed through the design process. In addition, grey in each figure indicates a non-creative version of each task, while black indicates a creative version.

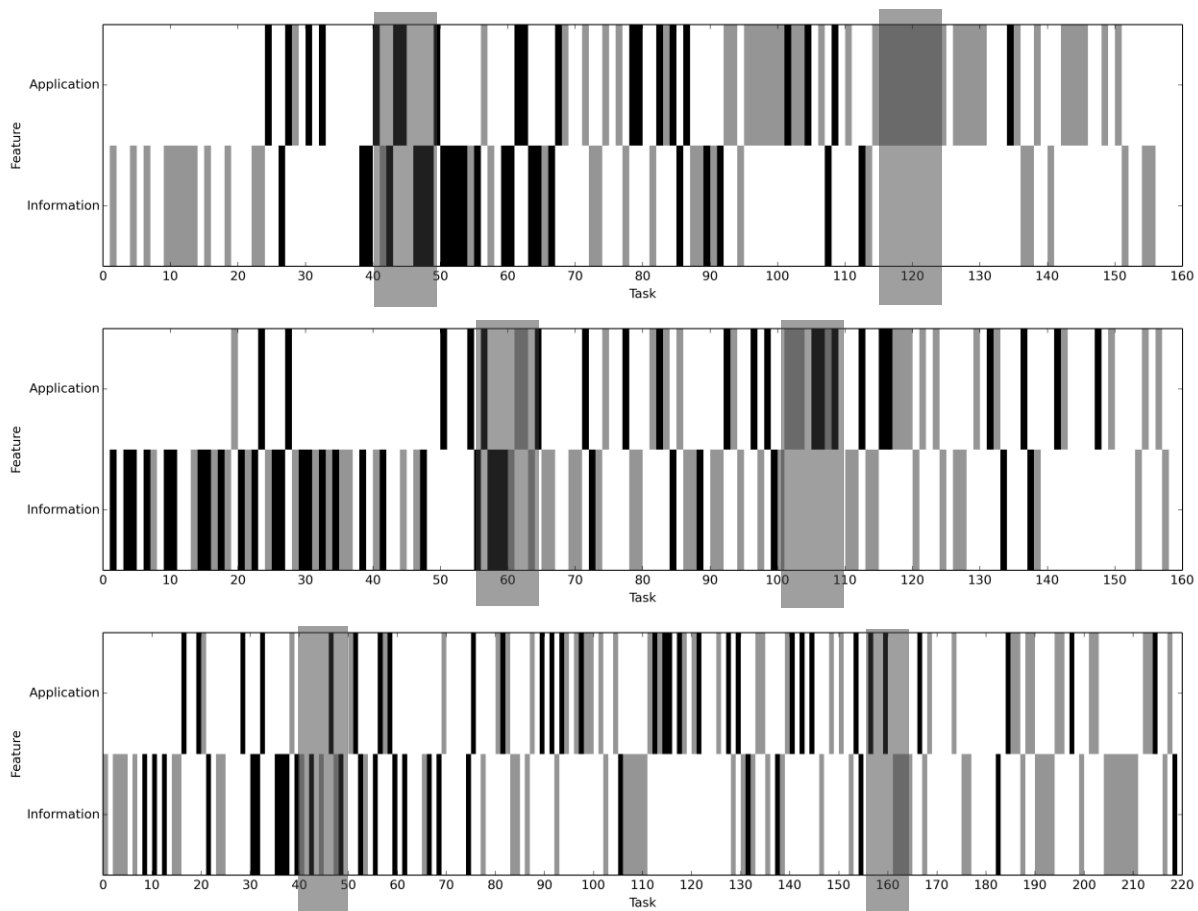


Figure 4: Task Types and Creative Tasks

From other research (Snider et al. 2013), quantities of such activity types are known to vary from a majority of information-type tasks (lower in each sub-figure) in early stages, to a majority of application-type (upper in each sub-figure) activities in later stages. Further, a majority of creativity in information-type tasks is thought to lead to higher novelty in solution, while a majority of creativity in application-type tasks is thought to lead to higher quality (Snider 2014). As such, there are potential actions to be taken by a manager based on the data when presented to them. For example, should a high amount of information-task creativity occur, suggesting a higher novelty output is likely, the manager may need to intervene. Particularly when novelty is likely to cause issue (such as in change propagation through complex systems or when cost constraints or time constraints are tight) the manager may wish to encourage non-creative behaviour. Alternatively, should creative behaviour in application tasks produce a higher quality result, an understanding of how much creative behaviour is occurring in such tasks may lead a manager to intervene if they deem necessary.

In each case, the knowledge of what is happening in the design process from the perspective of creativity gives the means for a manager to understand the process that is occurring in more detail, and promotes intervention and improvement of the process when progression is not as expected, or when opportunities for improvement arise.

4 DISCUSSION

Logbooks are a potentially interesting source of information on engineering projects because they contain information that is not found elsewhere (McAlpine et al. 2006; McAlpine et al. 2009) and being informal and personal in nature, are often the first type of information generated by engineers on a project. They therefore have good potential for 'advance warning' of issues, such as a failure to consider costs or risks, a lack of planning, or more positive things such as being the first written record of IP creation.

Because of this, we have sought to evaluate the utility of logbook as a source of information useful to those with an interest in project performance. We have shown that through analysis of logbooks, a manager may be aided in their understanding of: i) What is happening in terms of content creation; ii) The type of the tasks being (or not being) worked on; iii) How the work being completed leads to novel work; and iv) How the work completed leads to higher quality or more radical design, and compares to "normal".

However, it is quite clear that large differences between projects and engineers and their working styles demands a level of caution when interpreting a particular pattern or feature in the data in a prescribed way. There is definite individual difference between the engineers, which could be due to their own style other project influencers. For example, high levels of evaluation may be needed in a high novelty design processes, or could possibly be interpreted as the engineer being less expert. This means that logbook data cannot be used purely algorithmically. Critically, the person interpreting the final data must contextualise what they see and contrast against what they expect in each case.

4.1 Further Work

This preliminary study has shown the potential of engineering logbooks as a source of data to provide actionable insights into engineering projects. In order to gain a more complete picture of project activity, two types of further analysis could be undertaken: First, one could combine elements of both schemas. For example, indicators of creativity as indicated by Snider *et al.* (2013) schema, together with 'solving' activities related to the product and a decision could indicate a high potential for IP to have been created. Second, whilst logbooks are an important source of some types of information, it has been shown that different sources of information have different, often complimentary characteristics (McAlpine et al. 2009). It would therefore be of value to explore how logbooks interact with other sources of information, both informal and formal. For example, does a high number of logbook entries related to the product being design precede changes to digital files such as CAD models? - if not, why not?

It must also be recognised that this paper presents the results of a manual coding exercise, which would not scale to larger projects. There is still clearly a challenge around the automatic extraction of such information. Further, whilst this research did not seek to understand logbooks to inform the design of digital versions, this approach could also be used to inform the digitisation of the logbooking process from a different perspective - instead of seeking to support the engineer, the focus of the digitisation could be on supporting the project management process.

5 CONCLUSIONS

Engineering projects are large, complex and frequently distributed. Consequently it is difficult to understand and monitor whether the project is progressing in the desired or optimal manner. Even deciding what 'optimal' looks like for a project is problematic because of the difficulty in properly evaluating past projects. It is for these reasons that guidance on what makes successful projects is often generic and vague.

To address this we take a different approach, which aims to provide information to engineers and project managers on the status of a project, and past projects. This approach has so far seen various techniques applied to digital information such as emails, CAD files and reports. However, projects

also produce a large volume of informal information (one of the most important of which are paper-based engineering logbooks) which have been shown to contain information not found in more formal sources (McAlpine et al. 2009). We therefore sought to determine whether such logbooks could be used to provide insights that may contribute to a better understanding of project performance.

Three logbooks were coded in detail with two schemas focussing on content and context. The results show promise, with evidence that indicators of project performance, such as progression against plans, design focus, activity type, product novelty, and process stage may all be possible. Further, it is posited that richer insights could be elicited through further analysis, by both combining the two coding schemas used, or by combining the results with those from analysing other types of information, such as email and CAD files, to build a more complete picture of project performance.

ACKNOWLEDGMENTS

This work has been performed as part of the ESPRC Language of Collaborative Manufacture grant based at the University of Bristol and University of Bath. Supporting data can be requested from the authors, and access to supporting data will be granted subject to retrospective consent being requested and granted from the project participants.

REFERENCES

- Blessing, L.T.M. & Chakrabarti, A., 2009. *DRM, a Design Research Methodology*, New York: Springer.
- Cash, P. & Snider, C., 2014. Investigating design: A comparison of manifest and latent approaches. *Design Studies*, 35(5), pp.441–472.
- Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1), pp.37–46.
- Dorst, K., 2003. The problem of design problems. *Expertise in design*, pp.135–147.
- Dunnette, M.D., Campbell, J. & Jaastad, K., 1963. The effect of group participation on brainstorming effectiveness for 2 industrial samples. *Journal of applied psychology*, 47(1), p.30.
- Gross, M.D. & Do, E.Y., 1996. Demonstrating the electronic cocktail napkin: A paper-like interface for early design. In *Computer Human Interaction (CHI 96)*. Vancouver, BC.
- Guilford, J.P., 1956. The structure of intellect. *Psychological Bulletin*, 53(4), pp.267–293.
- Hayes, A.F. & Krippendorff, K., 2007. Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), pp.77–89.
- Katz, R., 1982. The Effects of Group Longevity on Project Communication and Performance. *Administrative Science Quarterly*, 27(1), pp.81–104.
- Klenke, K., 2008. *Qualitative research in the study of leadership*, Emerald.
- McAlpine, H. et al., 2006. An investigation into the use and content of the engineer's logbook. *Design Studies*, 27(4), pp.481–504.
- McAlpine, H., Hicks, B.J. & Culley, S.J., 2009. Comparing the information content of formal and informal design documents: lessons for more complete design records. *ICED09: International conference on engineering design*.
- Oehlberg, L., Lau, K. & Agogino, A.M., 2009. Tangible interactions in a digital age: Medium and graphic visualization in design journals. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, 23, pp.237–249.
- OGC, 2009. *Managing Successful Projects with PRINCE2*, London, UK.
- Pahl, G. & Beitz, W., 1984. *Engineering Design: A Systematic Approach*, London: Springer.
- Pinto, J.K. & Slevin, D.P., 1987. Critical factors in successful project implementation. *Engineering Management, IEEE Transactions on*, (1), pp.22–27.
- Pugh, S., 1990. *Total Design: integrated methods for successful product engineering*, Harlow: Prentice Hall.
- Simon, H.A., 1973. The structure of ill structured problems. *Artificial Intelligence*, 4(3-4), pp.181–201.
- Snider, C., 2014. *Characterising the Creative Behaviour of Designers within the Late-Stage Engineering Design Process*. University of Bath, UK.
- Snider, C. et al., 2015. Understanding Engineering Projects: An Integrated Vehicle Health Management Approach to Engineering Project Monitoring. In *ICED15: International Conference on Engineering Design*. Milan, Italy.
- Snider, C., Culley, S.J. & Dekoninck, E.A., 2013. Analysing creative behaviour in the later stage design process. *Design Studies*, 34(5), pp.543–574.
- Snider, C., Dekoninck, E.A. & Culley, S.J., 2014. The appearance of creative behavior in later stage design processes. *International Journal of Design Creativity and Innovation*, 2(1), pp.1–19.
- Wasiak, J. et al., 2010. Understanding engineering email: The development of a taxonomy for identifying and classifying engineering work. *Research in engineering design*, 21(1), pp.43–64.